



UvA-DARE (Digital Academic Repository)

Lured into listening

Engaging games as an alternative to reward-based crowdsourcing in music research

Honing, H.

DOI

[10.1027/2151-2604/a000474](https://doi.org/10.1027/2151-2604/a000474)

Publication date

2021

Document Version

Final published version

Published in

Zeitschrift für Psychologie

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

Honing, H. (2021). Lured into listening: Engaging games as an alternative to reward-based crowdsourcing in music research. *Zeitschrift für Psychologie*, 229(4), 266-268. <https://doi.org/10.1027/2151-2604/a000474>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.



Lured Into Listening

Engaging Games as an Alternative to Reward-Based Crowdsourcing in Music Research

Henkjan Honing

Music Cognition Group, Institute for Logic, Language and Computation, University of Amsterdam, The Netherlands

Abstract: This brief statement revisits some earlier observations on what makes web-based experiments, and especially citizen science using engaging games, an attractive alternative to laboratory-based setups. It suggests web-based experimenting to be a full-grown alternative to traditional laboratory-based experiments, especially in the field of music cognition, where sampling bias is a common problem and large amounts of empirical data are needed to characterize individual variability.

Keywords: web-based experiments, citizen science, methodology, music cognition, comparative musicology

For the field of music cognition, an important impact of the arrival of the Internet was that it offered researchers a much longed-for opportunity to do listening experiments outside the laboratory. Next to the advantages of versatility and the external validity of the experimental results, web-based experiments can – as became clear over the years –, attract a much larger, more diverse, and intrinsically motivated group of participants, as compared to the usual laboratory experiment (Germine et al., 2012; Honing, 2010; Honing & Ladinig, 2008; Honing & Reips, 2008). These three characteristics continue to play an important role in the domain of music cognition and contribute to a complete understanding of our capacity for music (Honing et al., 2015).

In this short statement, I will revisit some earlier observations on what makes web-based experiments, and especially citizen science using engaging games, an attractive alternative to laboratory-based setups, and will outline some promising directions for the near future.

On Standardization, Control, and Reward in Web-Based Experiments

While research in the cognitive sciences is increasingly using web-based setups (accelerated during the COVID-19 pandemic) and taking advantage of online crowdsourcing platforms, such as Amazon's Mechanical Turk (MTurk), for the collection of large amounts of empirical data, there is a continuing concern with the issue of *replicability* (Stewart et al., 2017) and the apparent lack of control one has in web-based as opposed to laboratory-based experiments. Where in the laboratory, most relevant factors,

including all technical issues (like presentation of the instructions, sound quality of the auditory stimuli, etc.), are under the control of the experimenter (they have a high *internal validity*), it is argued that web-based experiments lack this important foundation of experimental psychology (Bridges et al., 2020; Kendall, 2008; Mehler, 1999). These authors continue to worry about how the stimuli are presented and experienced at the user-end, something that appears to be beyond the control of the experimenter.

However, in contrast, it can be argued that web-based experiments have a much greater *external validity* than laboratory-based experiments. While experiments performed over the Internet sometimes lose some internal validity, in music cognition studies this might actually be desirable. It might be a better reflection of the listening environment of the participants, including its noisiness, use of low-quality headphones, and the like. In addition, it might be the *invariants* that participants are sensitive to (cf. Krantz, 2021, citing J. J. Gibson), not the technical variance.

Some authors even argue that experimental control and standardization should be seen as a cause of, rather than a cure for, poor reproducibility of experimental outcomes. As an example, Richter et al. (2010) showed that environmental standardization could contribute to spurious and conflicting findings in the literature. Their advice is, rather than to generate results that are most likely going to be reproducible in other laboratories, the strategies to standardize environmental conditions in an experiment should be minimized. In fact, both the technological and environmental variability introduced by web-based setups (and that is often contrasted with laboratory-based studies) might actually amount to experimental results with a much higher external validity than before.

And lastly, web-based experiments, because of their potential to obtain large amounts of empirical data, can reveal underlying perceptual and cognitive mechanisms that are not readily observed in the laboratory (see, e.g., Langlois et al., 2021, for a data-driven approach studying perception using crowdsourcing).

Toward a Larger, More Varied, and Motivated Participant Pool

While web-based experiments can be argued to have a greater external validity (as discussed above), they also can reach a potentially much larger and more varied participant pool (cf. Sheskin et al., 2020). Especially if the web-based experiments are designed as an engaging game and provide personalized feedback, we know that it can easily attract tens of thousands of dedicated participants and introduce considerable statistical power (Burgoyne et al., 2013; Mehr et al., 2019).

Much current research in cognitive science takes advantage of this scale and relies on paid participants recruited through crowdsourcing platforms like MTurk. This is estimated by Stewart et al. (2017) to be 50% of all recently published cognitive science research. However, the quality of such reward-based data cannot always be guaranteed (Buhrmester et al., 2011; Chmielewski & Kucker, 2020). For instance, it is possible that participants' motivation interacts with their performance in listening experiments (such as feeling obliged to listen for a long time). An alternative approach is to recruit participants by using *engaging listening games* (Honing, 2010), an approach that depends on the intrinsic motivation of the participants and hence will provide more valid data and less fraud or drop-out (Honing & Reips, 2008).

Still, one of the continuing challenges of web-based listening experiments is attracting a suitable participant group that is willing to seriously engage in online experiments. For this, several techniques are available that make such experiments attractive and intrinsically motivating (Aljanaki et al., 2014; Burgoyne et al., 2013), avoiding monetary rewards like MTurk (see above), course credit, or other motivations that could interfere with the quality of the responses.

Advantages of Intrinsically Motivating Games for Music Research

Engaging listening games allow one (as suggested by some case studies, see below) to probe music cognition across many different cultures, societies, and environments on an unprecedented scale. As such, it can avoid the sampling bias from which much music cognition research suffers (Jacoby et al., 2020). Furthermore, the scale of data collection allows one to map out the capacity for music (i.e., *musicality*) and its variability (Honing, 2018), based on the distribution of

certain musical traits (e.g., the ability to hear a regular beat in music; Bouwer et al., 2018). It can provide the large amounts of phenotypic data needed to search for correlations with variations at the genetic level (e.g., using genome-wide association scans) and the associated environmental variables (for a recent example, see Niarchou et al., 2021). All this will further encourage the exchange of knowledge and methodologies between the fields of music cognition, genetics, and cognitive biology (Gingras et al., 2015).

Some Trends for the Near Future

The main promise for web-based experiments in the field of music cognition lies, I think, in the development of intrinsically motivating games and the application of recent *citizen science* techniques that reveal in a natural way the behavior one is interested in. A successful citizen science experiment design is, of course, a challenge and might take more effort than a traditional laboratory experiment. But the few examples from the music domain that are currently available are quite promising (see, e.g., Harvard's themusiclab.org and the amsterdammusiclab.nl).

A recent example is "Hooked On Music," a citizen science project developed to uncover what makes music memorable. This game has been played 2 million times by nearly 200,000 participants in more than 200 countries. The game format is currently used for cross-cultural studies on music cognition and musical memory (Burgoyne et al., 2013; Honing, 2010). Of course, issues like privacy, confidentiality, reliability, and fraud continue to be a serious concern, but they are not fundamentally different from the issues that have to be dealt with for laboratory-based experiments (Garaizar & Reips, 2019; Honing & Reips, 2008).

In short: to secure reliability and validity (and avoid fraud or drop out) it seems wise to make a web-based experiment challenging and fun, not rewarding good answers, but simply participating, and make certain the participants feel involved (e.g., by giving personalized feedback). Overall, engaging games serve as an attractive alternative to reward-based crowdsourcing, and could attract, in principle, (1) a much larger, (2) more diverse, and (3) intrinsically motivated group of participants. These three characteristics continue to play an important role in the domain of music cognition and contribute to a complete understanding of our capacity for music (Honing et al., 2015).

References

- Aljanaki, A., Bountouridis, D., Burgoyne, J. A., van Balen, J., Wiering, F., Honing, H., & Veltkamp, R. (2014). Designing games with a purpose for data collection in music research. *Emotify and hooked: Two case studies*. In A. Aljanaki & A. De Gloria (Eds.), *Games and learning alliance* (pp. 29–40). Springer International Publishing.

- Bouwer, F. L., Burgoyne, J. A., Odiijk, D., Honing, H., & Grahn, J. A. (2018). What makes a rhythm complex? The influence of musical training and accent type on beat perception. *PLoS One*, 13(1), 1–26. <https://doi.org/10.1371/journal.pone.0190322>
- Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, 8, 1–29. <https://doi.org/10.7717/peerj.9414>
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1), 3–5. <https://doi.org/10.1177/1745691610393980>
- Burgoyne, J. A., Bountouridis, D., van Balen, J., & Honing, H. (2013). Hooked: A game for discovering what makes music catchy. In A. De Souza Britto, F. Gouyon, & S. Dixon (Eds.), *Proceedings of the International Society for Music Information Retrieval Conference* (pp. 245–250). Curitiba. <http://igitur-archive.library.uu.nl/math/2013-0904-200636/UUindex.html>
- Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science*, 11(4), 464–473. <https://doi.org/10.1177/1948550619875149>
- Garaizar, P., & Reips, U.-D. (2019). Best practices: Two Web-browser-based methods for stimulus presentation in behavioral experiments with high-resolution timing requirements. *Behavior Research Methods*, 51(3), 1441–1453. <https://doi.org/10.3758/s13428-018-1126-4>
- Germiné, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the web as good as the lab? Comparable performance from web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin and Review*, 19(5), 847–857. <https://doi.org/10.3758/s13423-012-0296-9>
- Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E. (2015). Defining the biological bases of individual differences in musicality. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1664), Article 20140092. <https://doi.org/10.1098/rstb.2014.0092>
- Honing, H. (2010). Lure (d) into listening: The potential of cognition-based music information retrieval. *Empirical Musicology Review*, 5(4), 121–126. <https://kb.osu.edu/dspace/handle/1811/48549>
- Honing, H. (Ed.). (2018). *The origins of musicality*. The MIT Press. <https://mitpress.mit.edu/books/origins-musicality>
- Honing, H., & Ladinig, O. (2008). The potential of the Internet for music perception research: A comment on lab-based versus Web-based studies. *Empirical Musicology Review*, 3(1), 4–7. <https://kb.osu.edu/dspace/handle/1811/31692>
- Honing, H., & Reips, U.-D. (2008). Web-based versus Lab-based studies: A response to Kendall (2008). *Empirical Musicology Review*, 3(2), 73–77.
- Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1664), Article 20140088. <https://doi.org/10.1098/rstb.2014.0088>
- Jacoby, N., Margulis, E. H., Clayton, M., Hannon, E., Honing, H., Iversen, J., Klein, T. R., Mehr, S. A., Pearson, L., Peretz, I., Perlman, M., Polak, R., Ravignani, A., Savage, P. E., Steingo, G., Stevens, C. J., Trainor, L., Trehub, S., Veal, M., & Wald-Fuhrmann, M. (2020). Cross-cultural work in music cognition: Challenges, insights and recommendations. *Music Perception*, 37(3), 185–195. <https://doi.org/10.1525/mp.2020.37.3.185>
- Kendall, R. (2008). Commentary on “The potential of the Internet for music perception research: A Comment on lab-based versus Web-based studies” by Honing & Ladinig. *Empirical Musicology Review*, 3(2), 8–10.
- Krantz, J. H. (2021). Ebbinghaus illusion: Relative size as a possible invariant under technically varied conditions? *Zeitschrift für Psychologie*, 229(4), 230–235. <https://doi.org/10.1027/2151-2604/a000467>
- Langlois, T. A., Jacoby, N., Suchow, J. W., & Griffiths, T. L. (2021). Serial reproduction reveals the geometry of visuospatial representations. *Proceedings of the National Academy of Sciences of the United States of America*, 118(13), 1–11. <https://doi.org/10.1073/pnas.2012938118>
- Mehler, J. (1999). Experiments carried out over the Internet. *Cognition*, 71, 187–189. [https://doi.org/10.1016/S0010-0277\(99\)0029-3](https://doi.org/10.1016/S0010-0277(99)0029-3)
- Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, 366(6468), Article eaax0868. <https://doi.org/10.1126/science.aax0868>
- Niarchou, M., Gustavson, D., Sathirapongsasuti, J. F., Anglada-Tort, M., Eising, E., Bell, E., McArthur, E., Straub, P., The 23andMe Research Team, McAuley, J. D., Capra, J. A., Ullén, F., Creanza, N., Mosing, M. A., Hinds, D., Davis, L. K., Jacoby, N. J., & Gordon, R. L. (2021). *Unravelling the genetic architecture of musical rhythm: A large-scale genome-wide association study of beat synchronization*. *BioRxiv*. <https://doi.org/10.1101/836197>
- Richter, S. H., Garner, J. P., Auer, C., Kunert, J., & Würbel, H. (2010). Systematic variation improves reproducibility of animal experiments. *Nature Methods*, 7(3), 167–168. <https://doi.org/10.1038/nmeth0310-167>
- Sheskin, M., Scott, K., Mills, C. M., Bergelson, E., Bonawitz, E., Spelke, E. S., Fei-Fei, L., Keil, F. C., Gweon, H., Tenenbaum, J. B., Jara-Ettinger, J., Adolph, K. E., Rhodes, M., Frank, M. C., Mehr, S. A., & Schulz, L. (2020). Online developmental science to foster innovation, access, and impact. *Trends in Cognitive Sciences*, 24(9), 675–678. <https://doi.org/10.1016/j.tics.2020.06.004>
- Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing samples in cognitive science. *Trends in Cognitive Sciences*, 21(10), 736–748. <https://doi.org/10.1016/j.tics.2017.06.007>

History

Received April 11, 2021

Revision received July 29, 2021

Accepted July 29, 2021

Published online December 17, 2021

Acknowledgments

I would like to thank J. Ashley Burgoyne, Bas Cornelissen, Samuel Mehr, and two anonymous reviewers for their feedback on an earlier version of this manuscript.

Funding

The author is supported by an 299 NWO Open Competition SSH grant (ref. 406.20.CW.002). Open access publication enabled by Hogrefe Publishing.

Henkjan Honing

Music Cognition Group (MCG)
Institute for Logic, Language and Computation (ILLC)
University of Amsterdam
P.O. Box 94242
1090 GE Amsterdam
The Netherlands
honing@uva.nl